Strategic Evolution of ESE Data Systems (SEEDS)

Briefing to ESISS Subcommittee
July 10, 2003

Presented By

Dr. Stephen Wharton
stephen.w.wharton@nasa.gov

SEEDS Study Team

Catherine Corlan, Kathy Fontaine, Vanessa Griffin,
Gail McConaughy, Ken McDonald, Karen Moe,
H. Ramapriyan, Richard Ullman, Stephen Wharton

SEEDS Web Page
http://eos.nasa.gov/seeds

Strategic Evolution of ESE Data Systems - Overview

SEEDS Mission

Establish evolution strategy and coordinating activities to assure continued effectiveness of ESE data management systems & services.

SEEDS Objectives

1. Ensure timely delivery of Earth Science information at an affordable cost.

2. Maximize availability and utility of ESE products.

3. Engage community on data management issues, objectives, and solutions.

4. Enable the development of flexible systems to readily accommodate evolving products & services.

**SEEDS Coordinating Responsibilities**

Support ESE Data Systems Evolution Planning
- Address key ESE data systems goals
- Coordinate data system evolution with implementing projects *
- Lead coordination/planning with national & international partners
- Support Transition to Measurement-Focused Paradigm
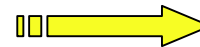- Sustain and Apply Cost Estimation Tool *

Sustain Community Involvement
- Conduct Community Workshops
- Support Four Working Groups *
- Brief Science Committees & Organizations
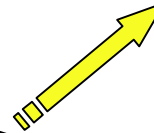
Foster Technology Evolution
- Provide Cost Estimation Tool *
- Carry Out IT Prototyping & Infusion
- Carry Out Software Reuse Initiatives

Sustain & Apply Unifying Framework of Core Standards & Guidelines
- Format & Interface Standards & Processes
- Levels of Service Guidelines
- Data Lifecycle Planning Procedures
- Metrics Planning & Reporting Guidelines
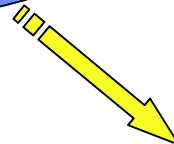
Support HQ Initiatives to Fund Distributed Providers of Products & Services
- Utilize SEEDS Paradigm
- Address Thematic Science Questions for Research, Education and Applications
- Support REASoN CAN Management *

SEEDS Status

➤ **Completed Study Recommendations:**
- ❑ Discussed draft recommendations at March 2003 workshop in Annapolis.
- ❑ We have community buy-in for the recommendations.
- ❑ Presented overview of recommendations to AA in March.
  - ❖ Received action from AA to develop plan for evolution of ESE data systems.
  - ❖ Working with ESDIS Project Manager to address this action (see slide next page).
- ❑ Incorporated feedback and published final recommendations July 3, 2003.

➤ **Supported REASoN CAN:**
- ❑ Contributed guidelines and selection criteria.
- ❑ Supported evaluation, selection, awards process and milestone negotiation.
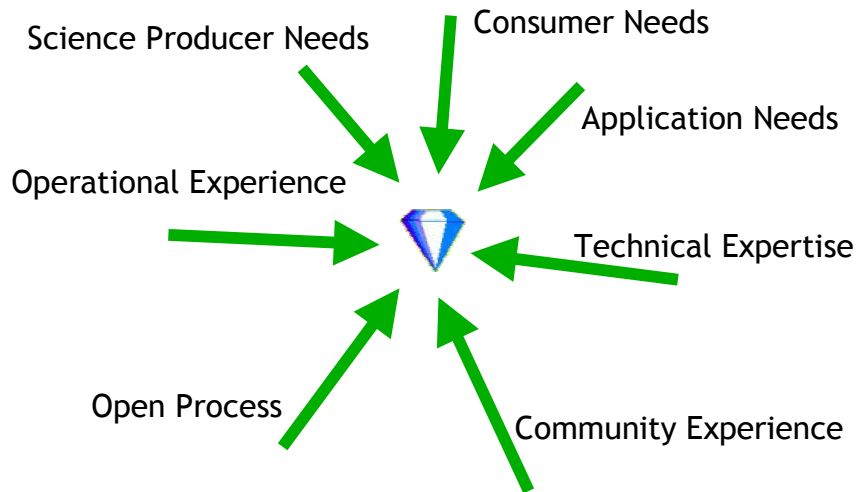
➤ **SEEDS planning activities to continue in FY2004:**
- ❑ Determination of the appropriate scope, responsibilities, and resources for SEEDS will be made as part of the ESE data systems planning action.
- ❑ Support working group participation by REASoN CAN awardees - standards, metrics, technology, reuse.
- ❑ Develop options and recommend approach and budget for SEEDS coordinating functions.

ESE Data Systems: Planning for the Future

- ➢ **Work in partnership with ESDIS to apply SEEDS principles and guidelines to plan the evolution of ESE data systems:**
  - ❑ SEEDS - focused on setting context for evolution of current data systems
  - ❑ ESDIS - focused on development, management and operation of Enterprise science data systems

- ➢ **Address ESE key data systems goals:**
  - ❑ Increase resources for higher level product generation - more science value
  - ❑ Increase community participation
  - ❑ Move to measurement based systems and away from mission-based systems
  - ❑ Increase utilization of smaller, distributed systems and reduce reliance on large, centralized systems
  - ❑ Enable the development of Climate Data Records
  - ❑ Closer collaboration between scientists and those planning and developing data systems

- ➢ **Define evolutionary path and project plan that:**
  - ❑ Meets ESE science and budgetary goals
  - ❑ Addresses data systems goals
  - ❑ Takes advantage of advances in technology
  - ❑ Utilizes existing assets to maximum ability
  - ❑ Is flexible enough to allow continued evolution

Background

- ❏ Standards Working Group - Overview
- ❏ Metrics Working Group - Status
- ❏ Technology Working Group - Status
- ❏ Reuse Working Group - Status
- ❏ Levels of Service and Cost Estimation Study - Status
- ❏ Data Lifecycle Study - Status
- ❏ SEEDS Study - Approach & Recommendations

Standards Working Group - Overview

Science Producer Needs

Consumer Needs

Application Needs

Operational Experience

Technical Expertise

Open Process

Community Experience

Results: Standards that are:

➢ **"Community" interoperable**

  Needed standards adopted, unnecessary standards avoided.

➢ **"Working edge"**

  Process-chosen standards proven to work.

➢ **Understood**

  Core ESE Standards help-desk supported.

➢ **Reliable**

  Science, applications and software developers trust that ESE data conformance.

Findings

➢ **There are no "one-size fits all" standards.**

➢ **Requirements for interchange differ from requirements for distribution.**

➢ **Implementation is as important as specification.**

➢ **Trust and standing of standards process is as important as technical factors.**

➢ **IETF process combines openness and decision making foundation** for fast evolving internet standards.

SEEDS Standards Actions

• Standards Process
  ➢ A community-involved standards adoption process acquires trusted standards on the "working edge".
• Direct Support to missions
  ➢ Assure that ESE standards are understood by affected ESE implementers (projects, missions, etc).
• Standards Participation & Development
  ➢ Maintain Enterprise voice in "leading edge" national and international standards.

Standards Working Group - Recommendations

➢ **Background:**
- ❑ Define standards for near-term missions (pending results from standards process).
- ❑ Recommend data format, metadata content, catalog interface, documentation.
- ❑ Distinguish between ESE network data interchange from distribution to end users.

➢ **Standards Management – Recommend that ESE:**
- ❑ Maintain format translators to distribute products in multiple formats.
- ❑ Upgrade interoperability capabilities (catalog, inventory, distribution).
- ❑ Plan for evolution of packaging requirements.
- ❑ Support ESE unique standards (development, maintenance, training, help desk).
- ❑ Support evolution of science data formats towards seamless operability.

➢ **Near-Term Mission Standards – ESE should require that:**
- ❑ Standard products be file based and use HDF or netCDF as an interchange format.
- ❑ Distribution format address user needs and convenience.
- ❑ Mission standard products:
  - ❖ Be further defined using profile.
  - ❖ Use ECS or ECHO data model for inventory metadata (pending ISO 19115 standard).
  - ❖ Documented using EOSDIS guide standard.
  - ❖ Use EOSDIS V0, Z39.50, or ECHO compatible order and distribution protocol

➢ **Long-Term Standards Processes – Recommend that ESE:**
- ❑ Adopt Internet Engineering Task Force Process tailored to meet specific ESE needs.
- ❑ Develop a strategy for facilitating ESE standards compliance across the enterprise, including the performance of standards support services, e.g. user support, training, tool development, etc.
- ❑ Encourage adoption of existing successful standards.
- ❑ Develop new standard if no existing viable candidates.

Standards Working Group - Status

- ➢ **FY 2003:**
  - ❑ 0.3 FTE civil servant; 2.0 FTE contractor
  - ❑ Revised SEEDS recommendations.
  - ❑ Prepared initial Working Group Charter; Standards Process; and RFC Instructions.

- ➢ **FY 2004:**
  - ❑ 0.3 FTE civil servant; 0.9 FTE contractor; participation from REASoN CAN
  - ❑ Describe proposed SEEDS standards process to REASON Standards panel.
  - ❑ Describe how other standards groups, particularly IETF, do their work.
  - ❑ Begin role/scope definition with OGI and CCSDS ISO Archiving Standards.
  - ❑ Review/revise working group recommendations and initial process documentation.

- ➢ **What will not be done:**
  - ❑ Help desk support for standards approved by the Reason Standards Panel (~1.2 FE).
  - ❑ Mission support for migration to new ESE standard (~1.0 FTE).
  - ❑ Funded participation in external standards bodies (~1.6 FTE).
  - ❑ Preparation of draft standards for consideration by the Standards working group (~1.0 FTE).
  - ❑ RFC editor and document handling (~0.75 FTE).
  - ❑ Web interface for submission of standards or the communication of decisions (~0.5 FTE).
  - ❑ Development of standards profiles by SPWG (~0.75 FTE).

- ➢ **Concerns:**
  - ❑ Standards working group does not yet have the charter and budget to be a fully leveraged Enterprise standards body:
    - ❖ Uncertain how (or if) the output of the process will be used.
    - ❖ Uncertain relationship with new missions.

Metrics Working Group - Status

- ➢ **FY 2003:**
  - ❑ 0.25 FTE civil servant; 0.6 FTE contractor
  - ❑ Completed draft SEEDS recommendations and incorporated community comments
  - ❑ Drafted working group charter
  - ❑ Surveyed candidate tools to support metrics collection

- ➢ **FY 2004:**
  - ❑ 0.25 FTE civil servant; 0.9 FTE contractor; participation from REASoN CAN
  - ❑ Refine metrics' definitions through MPAR Working Group with REASoN CAN participation
  - ❑ Identify common tools/capabilities needed

- ➢ **What will not be done:**
  - ❑ Develop/maintain common tools/capabilities for metrics collection (~1 FTE)

- ➢ **Concerns:**
  - ❑ May be a delay in facilitating metrics collection and reporting from REASoN organizations

- ➢ **FY 2003:**
  - ❑ 0.1 FTE civil servant; 0.3 FTE contractor
  - ❑ Completed draft SEEDS recommendations and incorporated community comments
  - ❑ Drafted working group charter
  - ❑ Surveyed candidate tech infusion strategies to model for SEEDS evolution

- ➢ **FY 2004:**
  - ❑ 0.1 FTE civil servant; 0.5 FTE contractor; participation from REASoN CAN
  - ❑ Present recommendations & findings; refine technology infusion process, identify key technologies for the near and mid-term (3 - 6 years) and make the case for future funding of infusion of technology

- ➢ **What will not be done:**
  - ❑ Development of a SEEDS vision short video for future data systems evolution (.25 FTE)
  - ❑ Development of open competition for '05 and beyond for continuing identification and initial implementation of key technologies for evolution, maturing technology readiness levels of TRL 6 to 9 (.75 FTE)

- ➢ **Concerns:**
  - ❑ Prior plans had assumed a 2005 start of competitively awarded technology infusion demonstration prototypes; delay affects ROI of technology utilization

Reuse Working Group Status

➢ **FY 2003:**
- ❑ 0.3 FTE civil servant; 1.4 FTE contractor
- ❑ Completed SEEDS recommendations, prepared initial Working Group Charter, Analyzing Intellectual Property Issues

➢ **FY 2004:**
- ❑ 0.1 FTE civil servant; 0.9 FTE contractor; participation from REASoN CAN
- ❑ Present recommendations, findings, ROI discussions with REASoN CAN and other working groups (e.g. SDWG)
- ❑ With working groups, define metadata and content description required by new mission developer to determine reuse of existing asset base alternatives
- ❑ Develop 'prototype" publishing mechanism leveraging existing infrastructure
- ❑ Overcome initial reuse impediment of determining "what is out there"

➢ **What will not be done:**
- ❑ Mission-critical prototype of actual reuse – that is, upgrade of existing asset for new system, publish in "repository" for further reuse (.75 FTE)
- ❑ Support for mission-success prototype within REASoN CAN (.25 FTE)
- ❑ Survey of future mission needs (.25 FTE)
- ❑ Development of open competition for '05 and beyond for continuing identification of existing reusable assets, upgrade of those assets via teaming between existing and future teams, and publishing of those assets in a repository (.75 FTE)

➢ **Concerns:**
- ❑ Prior plans had assumed starting in 2005 of 4/year competitively awarded upgrade of existing assets via teaming between existing and future teams, delay means delay in ROI of reuse
- ❑ No commitment of assets for future effort indicates continuing of ad-hoc reuse as in past

> ## FY 2003:
> - ❑ 0.3 FTE civil servant; 2.50 FTE contractor
> - ❑ Completed recommended levels of service
> - ❑ Completed prototype of the cost estimation tool

> ## FY 2004:
> - ❑ 0.2 FTE civil servant; 0.9 FTE contractor
> - ❑ Will likely not finish the cost estimation model before the end of FY 04.
> - ❑ Will not begin levels of service working group activity of LOS analysis, update, and coordination with cost estimation tool (1.25 FTE)
> - ❑ Will not perform on-going comparables database collection, refinement, and analysis (2.25 FTE)

- ➢ **FY 2003:**
  - ❑ 0.3 FTE civil servant; 0.5 FTE contractor
  - ❑ Completed recommendations with updates based on feedback received.

- ➢ **FY 2004:**
  - ❑ 0.3 FTE civil servant; 0.9 FTE contractor; participation from REASoN CAN.
  - ❑ Form Data Lifecycle Working Group (voluntary effort, no funding), and develop charter.
  - ❑ Generate data lifecycle issues list and recommend resolutions.
  - ❑ Draft data lifecycle requirements language for inclusion into NASA ESE solicitations.

➢ **Formed seven study teams:**
- ❑ Cost Estimation and Levels of Service
- ❑ Standards and Interfaces for Near Term Missions
- ❑ Standards and Interfaces for Future Missions
- ❑ Life Cycle Data Management for Long Term Archive
- ❑ Software Reuse and Reference Architecture
- ❑ Technology Needs and Infusion Planning
- ❑ Metrics Planning and Reporting

➢ **Study Team Approach:**
- ❑ Survey existing practices, capabilities, and lessons learned.
- ❑ Present and discuss plans, progress, recommendations at SEEDS Public Workshops.
- ❑ Refine survey, questions, recommendations in response to community feedback.
- ❑ Issue recommendations.

➢ **Provided multiple opportunities for community involvement & comment:**
- ❑ Community engaged as participants and consultants on seven study teams.
- ❑ Held 3 public workshops (University of Maryland, San Diego, Annapolis).
- ❑ Presented SEEDS plans and status to NASA advisory committees.

➢ **Published draft recommendations on web:**
- ❑ SEEDS Web Page http://eos.nasa.gov/seeds
- ❑ Main body of recommendations organized into 7 chapters (74 pages).
- ❑ Background information provided in appendix (300 pages).

SEEDS Recommendations

- ➢ **Levels of Service:**
  - ❑ Identifies minimum and recommended levels of service for core data sets and services.
  - ❑ Recommends that ESE adopt levels of service (and establish working group).

- ➢ **Standards:**
  - ❑ Identifies science data, metadata, and interoperability standards for consideration.
  - ❑ Describes process for SEEDS to develop, adopt, and evolve standards and interfaces.
  - ❑ Recommends that ESE adopt the standards process (and establish working group).

- ➢ **Cost Estimation & Benchmarking:**
  - ❑ Describes cost-by-analogy model to estimate cost of ESE data management services.
  - ❑ Recommends that ESE adopt model and acquire cost information funded providers.

- ➢ **Data Lifecycle Planning:**
  - ❑ Describes approach to safely migrate products from generation thru long term archive.
  - ❑ Recommends that ESE adopt the approach (need for a working group is TBD).

- ➢ **Reuse Assessment:**
  - ❑ Defines approach for a SEEDS investment in software reuse.
  - ❑ Recommends that ESE adopt the reuse approach (and establish working group).

- ➢ **Technology Prototyping & Infusion:**
  - ❑ Describes process to identify needs and infuse technology investments into ESE provider network.
  - ❑ Recommends that ESE adopt process, define SEEDS technology vision (and establish working group).

- ➢ **Metrics Planning and Reporting:**
  - ❑ Defines metrics and reporting requirements for the participants in ESE Data Management Activities.
  - ❑ Recommends that ESE adopt the proposed metrics & requirements (and establish working group).